Proceedings of the 2nd International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/47/20230365

Study on the Applicability of LSTM for Predicting Stock Price when Facing Extreme Events

Dingju Dong[1,a,*]

[1]*Business Administration, Northeastern University, No.195, Chuangxin Street, Shenyang, China*
a. 20211351@stu.neu.edu.cn
**corresponding author*

Abstract: The analysis of stock price fluctuations holds considerable significance in the field of economics, particularly given the present environment characterized by unpredictability and rapid changes. Previously, the long short-term memory (LSTM) model has been employed effectively in addressing time series problems, including stock market forecasting. However, in the current dynamic landscape, the ability of LSTM to adapt to volatile conditions and provide accurate predictions is an area that merits further investigation. This study gathers stock data from prominent and representative companies, namely Apple, Google, Amazon, and Microsoft, spanning from January 2012 to March 2023. Specifically, two significant events are examined: the impact of the Covid-19 outbreak on the US stock market on February 26, 2020, and the Russia-Ukraine conflict occurring on February 26, 2022. By dividing the stock data surrounding these events into training and test sets, this research aims to evaluate the differential performance of LSTM in scenarios where it possesses no prior knowledge of these events versus situations where it has already assimilated the influence exerted by them.

Keywords: LSTM, stock prediction, Covid-19, machine learning

1. Introduction

The trend of how the stock price changes holds significant importance within the field of economics [1]. That’s because the stock price can directly reflect the situation of the big companies and these big companies will undoubtedly have an effect on other smaller companies and finally it will influence everyone. So, the stock actually not only affects the investor, but it will affect the life of all the people. However, stock prices are affected by various internal and external factors, such as domestic economic environment, political problem, international situation, rapid development of technology and stock market operation [2]. Also, as a financial time-series, the characteristics of random walk can be easily found in stock data [1], which makes it rather difficult to make prediction on stock. Thus, the forecasting for stock prices is an important and challenging issue, which deserves much attention.

The Recurrent Neural Network (RNN) has emerged as a suitable model for analyzing time series data, particularly in the context of stock prices. While RNNs possess the ability to capture the sequential context of the data, they often face challenges with long-term dependencies. To address this issue, scholars have introduced the Long Short-Term Memory (LSTM) model, which was developed by Schmidhuber et al. in 1997 to mitigate problems related to the gradient explosion and

DOI: 10.54254/2754-1169/47/20230365

gradient disappearance in traditional RNNs [3]. The newly added Gating Mechanism in each unit of LSTM can make timely adjustments to long-term memory based on historical information and current input. It has been widely used in many areas like written number pattern recognition, gesture recognition and speech recognition. And it can make relatively more accurate forecasting compared to other neural network models as it has its own memory, which makes it become very popular in the solution of series problem. In recent years, it has also been adopted in the field of stock market forecasting and generally it performs well in the forecast [1].

The exclusion of extreme events, such as the COVID-19 pandemic or geopolitical conflicts like the Russia-Ukraine Conflict, from the training and testing of LSTM models raises important considerations regarding their ability to effectively capture and predict stock price trends during such exceptional circumstances. The performance of LSTM in these situations, as well as its capacity to discern and forecast stock price trends, present significant subjects for investigation. Furthermore, exploring methods to enhance the LSTM model's performance when confronted with unprecedented events is a topic of notable scholarly interest, holding the potential to advance the field of stock market prediction.

So, this paper mainly studies the above problems, and the primary thing is to collect the data about the stock's price in extreme situations. Specifically, data pertaining to renowned companies such as Apple, Google, and Microsoft from January 2012 to March 2023 were obtained to conduct an investigation into the aforementioned topics. This period of time includes the period when covid-19 spread, and Russia-Ukraine Conflict occurred. This data is representative of the stock market and can reflect this extreme influence. Using them to train the LSTM model can help draw a conclusion about the question mentioned. To make model comparison, some common metrics like Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) serve as relevant evaluation indices [4].

2. Method

2.1. Dataset Description and Preprocessing

The datasets utilized in this research encompass a time span of four representative companies: Apple, Google, Amazon and Microsoft ranging from January 1, 2012, to March 30, 2023, and were all obtained from Yahoo Finance. Specifically, the dataset includes 6 features: open, close, high, low, adj close and volume of the stock. To visualize the data, here is an example of the following Figure 1 and Figure 2 that represent the data of Apple and Google's stock.

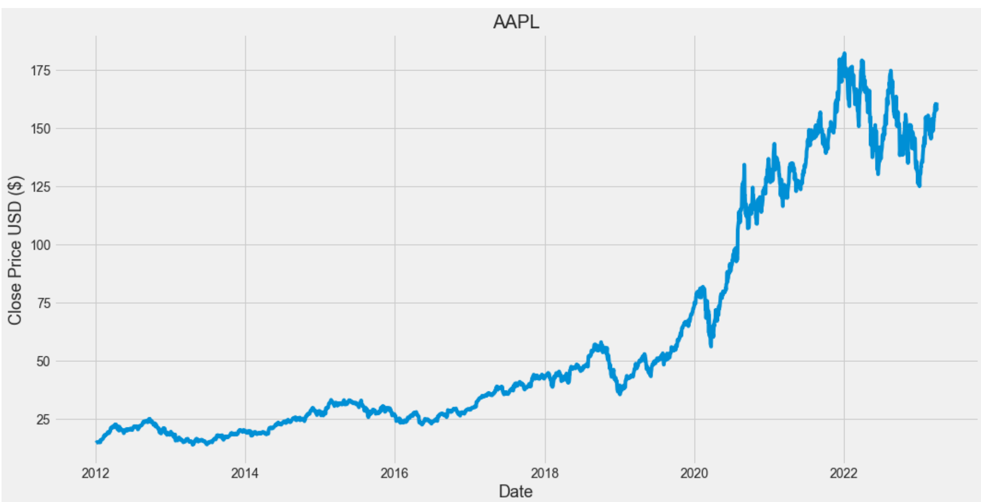


Figure 1: The closing price of apple stock (Photo/Picture credit: Original).

DOI: 10.54254/2754-1169/47/20230365

Figure 2: The closing price of google stock (Photo/Picture credit: Original).

The datasets employed in this study exhibit favorable characteristics as they contain no missing values, thereby obviating the need for imputation techniques. And to eliminate the impact of dimensionality, this paper used the technology called min max normalization to scale the data. Dissimilar to the common stock prediction model which always splits the data in the way of using 90% or 95% of all data as the training data and the rest will be treated as the test data, this paper splits the data in an unusual way. It focuses on two dates to split the data into training data and test data: February 26, 2020, and February 26, 2022. That is the date when Covid-19 cause the decline of the stock market of US and the date when the conflict between Russia and Ukraine occurred. These events can be considered as the extreme events this paper is studying about. Specifically, this paper uses the dates around these two dates to split the data, not the precise date, like 1 month before the events and 3 or 6 months after the events. That's because usually these events affect the stock market gradually, not so suddenly. Like before the Covid-19 hit the stock market in the US, it had spread for some time in US, and this already made some influence. So using the date 1 month before the date instead of the specific date would be better. Splitting data in this way, it can be observed that how LSTM will perform when it learns the influence of the extreme events or not.

2.2. LSTM Model

Neural Networks has been demonstrated in many tasks in the last decade [5-7]. Thereinto, to deal with the persistent challenges of gradient explosion and gradient vanishing in traditional Recurrent Neural Networks (RNN), Schmidhuber build a new neural network model in 1997 and called it Long Short-Term Memory (LSTM) [8]. It works well in the area of solving time-series problems. After it was carried out in 1997, people have used it to deal with problems like written number pattern recognition, text analysis, gesture recognition and speech recognition and LSTM can make relatively accurate forecasting [9].

A basic LSTM structure is shown in Figure 3 [10]. It's a breakthrough on LSTM compared to traditional RNN that it imports cell state, and it uses these three kinds of gates to remain and control the information: input gate, forget gate and output gate. All of this makes LSTM 'remember' much more information than traditional RNN model and improves the ability of neural network model on solving time-series problem. It's the reason that LSTM can be popularly adopted in many fields as there are always some time-series problems that need to be solved. And certainly, the forecasting on stock is typical time-series problem as people usually predict the trend of the stock with its previous price. It's easy to understand that if a stock remains 100$ for a long pass time, its price will possibly remain 100$ tomorrow.

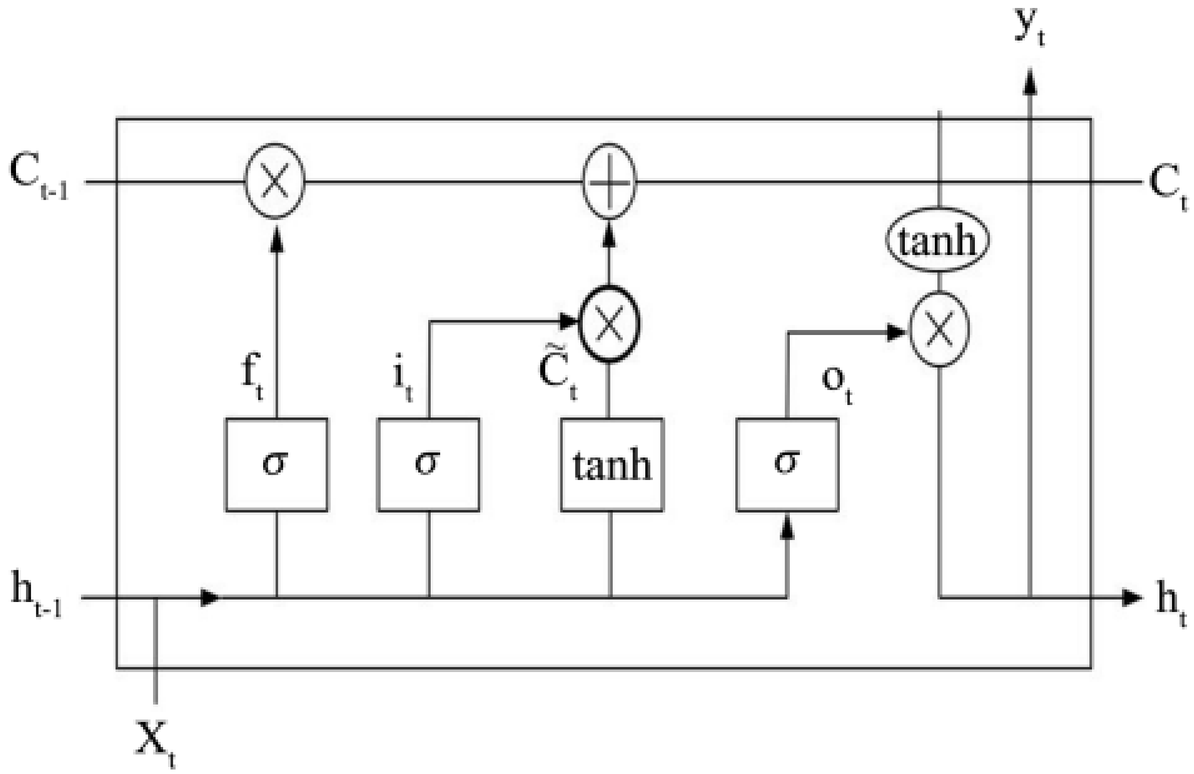


Figure 3: Cyclic unit structure of LSTM network [10].

The calculation process of LSTM at the moment of t is given as follows:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right), \tag{1}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \tag{2}$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \tag{3}$$

$$C_t = f_t * C_{t-1} + i_t * \widetilde{C}_t, \tag{4}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t * \tanh(C_t) \tag{6}$$

In these formulas, x_t is the input value of the current time, h_{t-1} is the output value of the last moment,$\tilde{C}_t$ is the candidate value of the memory cell in t moment,W_f is the weight of the forget gate,W_i is the weight of the input gate,W_c is the weight of the candidate input gate,W_o is the weight of the output gate, and b is the bias of the corresponding gate. The state of the memory cell is adjusted

DOI: 10.54254/2754-1169/47/20230365

by both of the input gate and forget gate and the final output of LSTM cell is decided by several functions.

2.3. Implementation Details

In this study, the LSTM model is implemented utilizing the Keras framework. This paper uses six features: open price, highest price, lowest price, close price, volume, adj close price of the stock data as the input. The data of 45 last days is used as the x_train and the next day, also can be called the 46 day, is recognized as the y_train. The optimizer is Adam optimizer and loss function chooses Mean Squared Error (MSE) and epochs equals to 10. Finally, Root Mean Squared Error (RMSE) Mean Absolute Error (MAE) are used to evaluate the performance of the model. The calculation process of these metrics is given as follows:

$$MSE = \frac{1}{n} * \sum_{i=1}^{n} \left(\widehat{X_i} - X_i\right)^2 \tag{7}$$

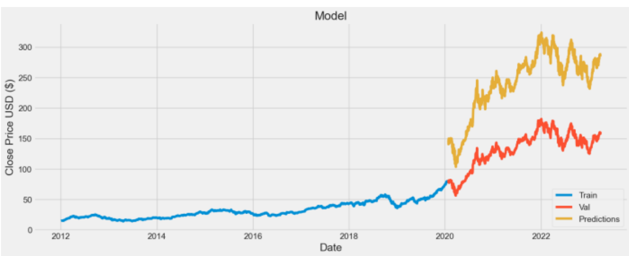
$$RMSE = \sqrt{\frac{1}{n} * \sum_{i=1}^{n} \left(\widehat{X_i} - X_i\right)^2} \tag{8}$$

$$MAE = \frac{1}{n} * \sum_{i=1}^{n} \left|\widehat{X_i} - X_i\right| \tag{9}$$

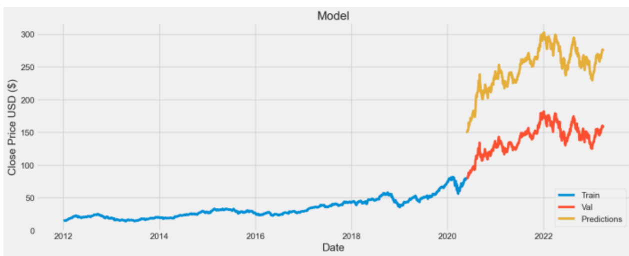
In these formulas, $\widehat{X_i}$ represents the estimated value of the prediction and X_i represents the true value. And in this paper, all the calculation of these metrics is done after scaling the data using min max scaler.

3. Results and Discussion

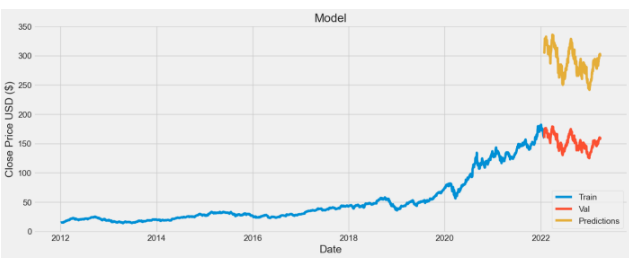
Figure 4, Figure 5, Figure 6 and Figure 7 present some examples of part of the predicted performance in different stocks based with different splits of data on the LSTM. Additionally, the Table 1 is the summary of all the prediction using LSTM with different training data.



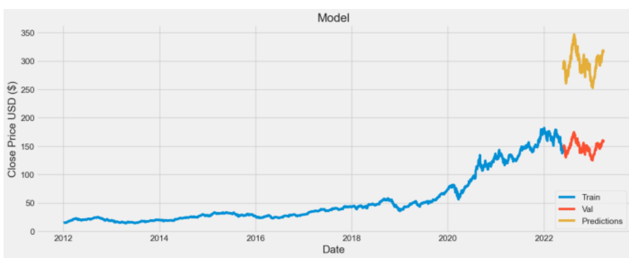
(a) 2020.1.26 to split data



(b) 2020.5.26 to split data



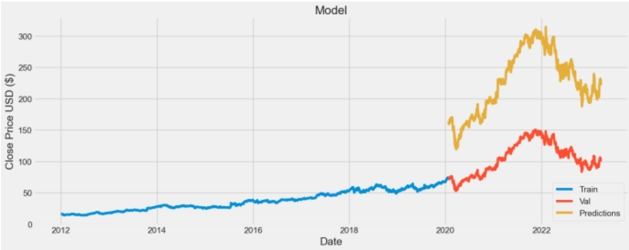
(c) 2022.1.26 to split data



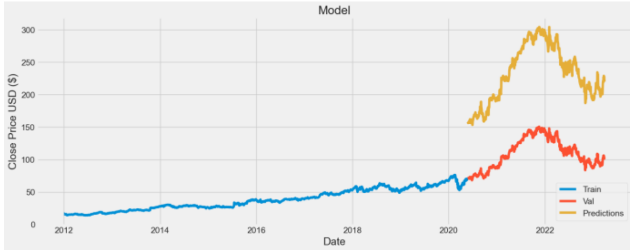
(d) 2022.5.26 to split data

Figure 4: Predictions on Apple's stock (Photo/Picture credit: Original).

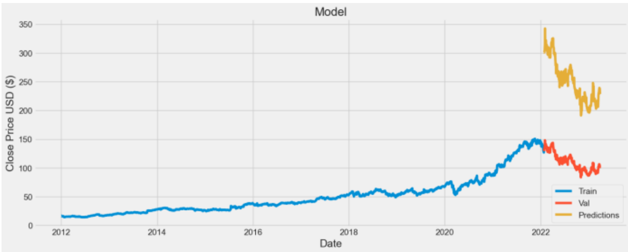
DOI: 10.54254/2754-1169/47/20230365



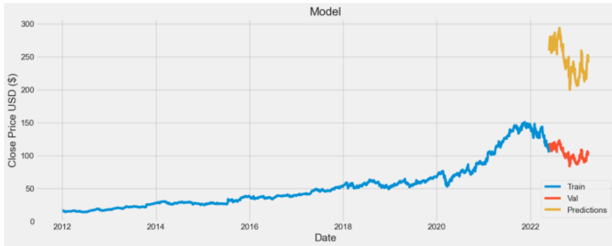
(a) 2020.1.26 to split data



(b) 2020.5.26 to split data

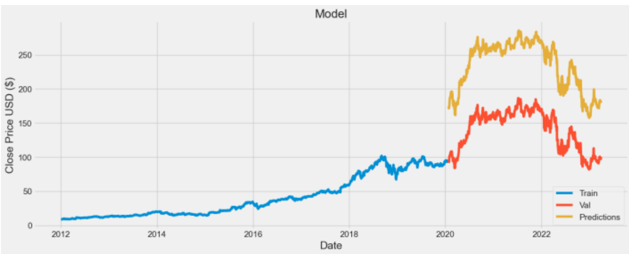


(c) 2022.1.26 to split data

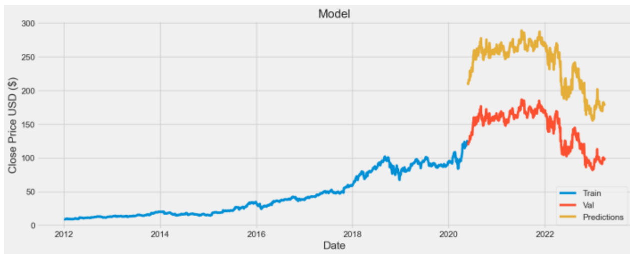


(d) 2022.5.26 to split data

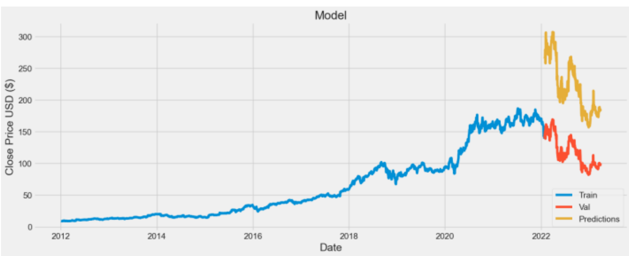
Figure 5: Predictions on Google's stock (Photo/Picture credit: Original).



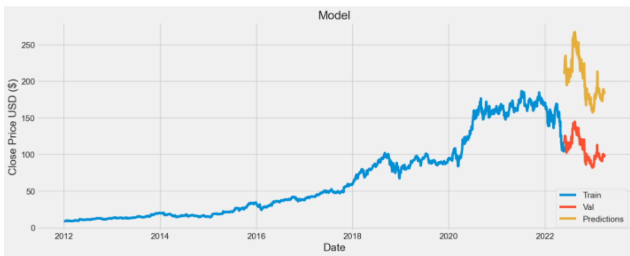
(a) 2020.1.26 to split data



(b) 2020.5.26 to split data



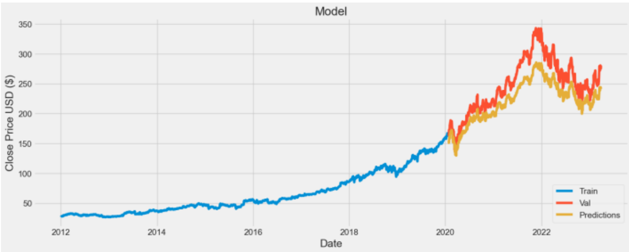
(c) 2022.1.26 to split data



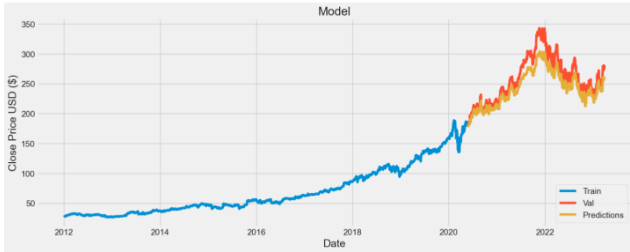
(d) 2022.5.26 to split data

Figure 6: Predictions on Amazon's stock (Photo/Picture credit: Original).

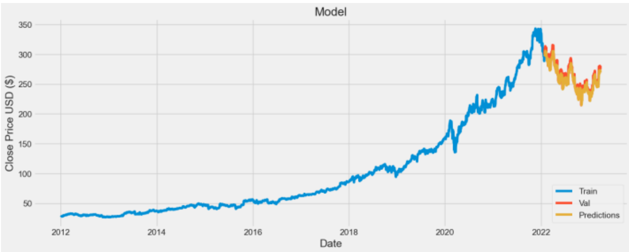
DOI: 10.54254/2754-1169/47/20230365



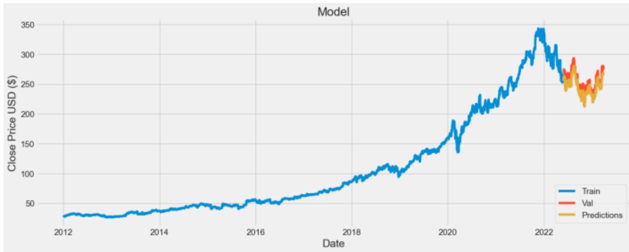
(a) 2020.1.26 to split data



(b) 2020.5.26 to split data



(c) 2022.1.26 to split data



(d) 2022.5.26 to split data

Figure 7: Predictions on Microsoft's stock (Photo/Picture credit: Original).

Because of the randomness in LSTM, some predictions on some stocks seem not so good. But this paper mainly studies the difference of the performance of LSTM when it learns the impact caused by extreme events. That's to say the change in the prediction when different data is provided is more important. The bias caused by the randomness in LSTM is acceptable. And to observe these changes directly, the following table makes a summary.

Table 1: The summary of all the predictions using LSTM with different training data.

The date to split data	**APPLE**		**GOOGLE**		**AMAZON**		**MICROSOFT**	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
2020.1.26	0.2401	0.1871	0.2477	0.2025	0.2139	0.1801	0.1935	0.1556
2020.5.26	0.1726	0.1372	0.2217	0.1814	0.2106	0.1773	0.1545	0.1242
Decrease(compare to 2020.1.26)	28.1%	26.67%	10.50%	10.42%	1.54%	1.55%	20.16%	20.18%
2020.8.26	0.1395	0.1129	0.2083	0.1701	0.2197	0.1704	0.1490	0.1195
Decrease(compare to 2020.1.26)	41.90%	39.66%	15.91%	16.00%	-2.71%	5.39%	23.00%	23.20%
2022.1.26	0.1007	0.0815	0.1619	0.1300	0.1852	0.1492	0.0993	0.0803
2022.5.26	0.0990	0.0795	0.1110	0.0901	0.1275	0.1023	0.0726	0.0584
Decrease(compare to 2022.1.26)	1.69%	2.45%	31.44%	30.85%	31.16%	31.43%	26.89%	27.27%
2022.8.26	0.0731	0.0583	0.0723	0.0578	0.1103	0.0857	0.0591	0.0471
Decrease(compare to 2022.1.26)	27.41%	28.47%	55.34%	55.54%	40.44%	42.56%	40.48%	41.34%

From Figure 8, the result can be shown more directly.

DOI: 10.54254/2754-1169/47/20230365

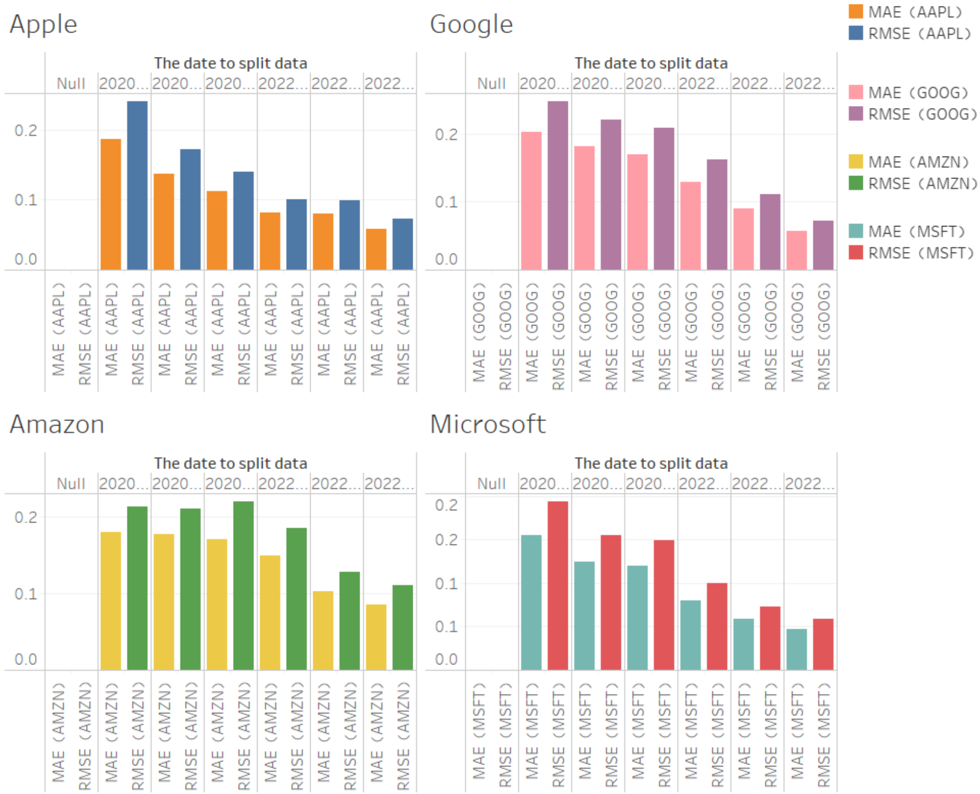


Figure 8: The pattern of the change of MAE and RMSE of different companies (Photo/Picture credit: Original).

The results indicate a notable trend: as the date used for splitting the data into training and test sets increases, both RMSE and MAE exhibit a decreasing pattern. This observation suggests an improvement in the performance of the LSTM model over time. And specifically, the model initially struggles to capture and incorporate the influence of significant events such as the Covid-19 pandemic or conflicts, resulting in relatively poorer predictions. Then if the data about 3 months after these events happened is provided, both RMSE and MAE can decrease about 20%, but there are some situations where the decrease is only 1-2%. When the data about 6 months later is provided, it's observed that most of the decrease on RMSE and MAE can reach up to 25% and half of them can reach more than 40%. In a few situations the decrease will not be so obvious.

Generally, the result shows though LSTM cannot give out an accurate prediction when some big events suddenly happen, it can perform much better after learning these events for some time. Furthermore, it means that LSTM has good resilience when facing some extreme events. LSTM can still be used in stock prediction in nowadays unexpected and changeful environments.

4. Conclusion

This paper focuses on the performance of LSTM model on predicting the stock's price when facing some extreme events like Covid-19 or Russia-Ukraine conflict. This paper selects four representative companies as an example and focuses on two dates: Feb 26,2020 and Feb 26,2022. The former date is the date when Covid-19 made big influence on the stock market of US and the latter is the day when Russia-Ukraine conflict started. To observe how LSTM performs differently when it learns the influence caused by these events, this paper chooses several dates around these two dates to split the data, like Jan 26,2020, May 26, 2020, and Aug 26, 2020. The result shows that LSTM model performs

DOI: 10.54254/2754-1169/47/20230365

better as newer data is provided. If data about 3 months later is provided, LSTM can generally perform better as both RMSE and MAE of most companies can decrease about 20%, and it can give out a far more precise prediction with the decrease of half of the companies can reach up to 40% when it gets data about 6 months after the events. To sum up, LSTM has good resilience when facing some extreme events, if the training data about these events is given, it can perform well.

References

[1] *Qiao, R., (2019) Stock prediction model based on neural network, Operations Research and Management Science, vol. 28, no. 10, pp. 132–140.*

[2] *Lu, W., et al. (2020) A CNN-LSTM-based model to forecast stock prices. Complexity, 2020: 1-10.*

[3] *Hochreiter, S., et al. (1997) Long short-term memory, MIT Press, vol. 9, no. 8, pp. 1735–1780.*

[4] *Prasannavenkatesan, T. (2020) Probable Forecasting of Epidemic COVID-19 in Using COCUDE Model for the State of Tamilnadu, India, Research Square Platform LLC*

[5] *Wang, F., Tax, D. M. J. (2016) Survey on the attention based RNN model and its applications in computer vision. arXiv preprint arXiv:1601.06823.*

[6] *Yu, Q., Wang, J., Jin, Z., et al. (2022) Pose-guided matching based on deep learning for assessing quality of action on rehabilitation training. Biomedical Signal Processing and Control, 72: 103323.*

[7] *Patel, J. L., Goyal, R. K. (2007) Applications of artificial neural networks in medical science. Current clinical pharmacology, 2(3): 217-226.*

[8] *Ta, V. D., Liu, C. M., Tadesse, D. A. (2020) Portfolio optimization-based stock prediction using long-short term memory network in quantitative trading. Applied Sciences, 10(2): 437.*

[9] *Gupta, N., et al. (2019) Integration of textual cues for fifine grained image captioning using deep CNN and LSTM, Neural Computing and Applications, vol. 12, pp. 1–10.*

[10] *Wang, X., et al. (2022) Research on Stock Forecasting Based on LSTM-ARIMA Model Chengdu Jincheng University, Chengdu Sichuan, Advances in Social Sciences, 11(7), 2843-2856.*